

¹ **A dynamic approach to addressing observation-minus-forecast**
² **mean differences in a land surface skin temperature data**
³ **assimilation system**

⁴ CLARA DRAPER^{1,2}, * ROLF REICHLE¹, GABRIELLE DE LANNOY^{1,2}

1. Global Modeling and Assimilation Office, NASA GSFC, Greenbelt, MD, USA

2. Universities Space Research Association, Columbia, MD, USA

⁵ BENJAMIN SCARINO³

3. Science Systems and Applications, Inc., Hampton, VA, USA

**Corresponding author address:* Clara Draper, Global Modeling and Assimilation Office, NASA GSFC, Code 610.1, Greenbelt, MD, 20771, USA.
E-mail: clara.draper@nasa.gov

ABSTRACT

7 In land data assimilation, bias in the observation-minus-forecast (O-F) residuals is typically
8 removed from the observations prior to assimilation by rescaling the observations to have
9 the same long-term mean (and higher-order moments) as the corresponding model fore-
10 casts. Such observation rescaling approaches require a long record of observed and forecast
11 estimates, and an assumption that the O-F mean differences are stationary. A two-stage
12 observation bias and state estimation filter is presented, as an alternative to observation
13 rescaling that does not require a long data record or assume stationary O-F mean differ-
14 ences. The two-stage filter removes dynamic (nonstationary) estimates of the seasonal scale
15 O-F mean difference from the assimilated observations, allowing the assimilation to correct
16 the model for synoptic-scale errors without adverse effects from observation biases. The
17 two-stage filter is demonstrated by assimilating geostationary skin temperature (T_{skin}) ob-
18 servations into the Catchment land surface model. Global maps of the O-F mean differences
19 are presented, and the two-stage filter is evaluated for one year over the Americas. The two-
20 stage filter effectively removed the T_{skin} O-F mean differences, for example the GOES-West
21 O-F mean difference at 21:00 UTC was reduced from 5.1 K for a bias-blind assimilation to 0.3
22 K. Compared to independent in situ and remotely sensed T_{skin} observations, the two-stage
23 assimilation reduced the unbiased Root Mean Square Difference (ubRMSD) of the modeled
24 T_{skin} by 10% of the open-loop values.

25 1. Introduction

26 Within the context of data assimilation, ‘bias’ refers to errors in modeled or observed
27 variables that persist over time and/or space. Standard ‘bias-blind’ data assimilation meth-
28 ods are based on the assumption that neither the forecast model nor the observations are
29 biased, and these methods will produce suboptimal output in the presence of bias (Dee and
30 Da Silva 1998). Unfortunately, the forecast models and observation data sets used in Earth
31 system applications, including for the land surface, typically are biased (Dee and Todling
32 2000; Reichle et al. 2004). Observation biases can arise from errors in the observing in-
33 strument and its calibration, the observation operator, or the retrieval model, as well as
34 representativity errors between the observed state variables and their modeled counterparts.
35 Likewise, forecast biases can arise from errors in the forecast model structure, parameters,
36 initial conditions, and forcing.

37 Ideally, the cause of observation and forecast biases should be diagnosed and treated at
38 the source. Where this is not possible, these biases can also be addressed in data assimilation
39 by applying an observation bias correction prior to assimilation (e.g., Harris and Kelly, 2001)
40 or by using a ‘bias-aware’ assimilation system explicitly designed to correct either observation
41 biases (e.g., Auligné et al. 2007; Fertig et al. 2009) or forecast biases (e.g., Dee and Todling
42 2000; Keppenne et al. 2005). Bias correction methods require that the bias be observable
43 (Dee and Da Silva 1998), and the ocean and atmosphere examples cited above measure the
44 biases against confident estimates of the true mean state, typically obtained with reference to
45 point-based observations (e.g., ocean buoys, radiosondes). However, the land surface is much
46 more heterogeneous than the ocean and atmosphere, and point-based in situ observations
47 are in general not representative of the coarse resolution states estimated by remote sensors
48 and land surface models (Crow et al. 2012). Consequently, for large domains the true mean
49 land surface states are unknown, since there are large systematic differences between the
50 mean (and variance) of different observed and modeled land surface data sets, none of which
51 can in general be identified as having statistics representative of the true state (Reichle et al.

52 2004).

53 Since observation and forecast biases cannot be observed for land surface states, it is
54 standard practice to remove the systematic differences between the observed and forecast
55 estimates from land data assimilation, usually by rescaling the observations to be consis-
56 tent with the long-term mean (and variance, and sometimes higher order moments) of the
57 forecasts (e.g., Reichle and Koster 2004; Drusch et al. 2005; Scipal et al. 2008; Crow et al.
58 2011). This prevents the systematic differences from adversely impacting the model state,
59 while satisfying the minimum criterion for optimal bias-blind data assimilation that there be
60 no difference between the mean values of the observed and forecast estimates. The assimi-
61 lation can then correct the model for random errors developing during each forecast, where
62 ‘random errors’ are errors persisting over time scales much shorter than the assumed bias
63 time scale. Data assimilation with observation rescaling has been shown to yield land surface
64 estimates that are superior to modeled or observed estimates alone (Slater and Clark 2006;
65 Reichle et al. 2007; Ghent et al. 2010; Crow et al. 2011; Draper et al. 2012; De Lannoy et al.
66 2012; de Rosnay et al. 2013). This rescaling approach is often referred to as ‘observation
67 bias correction’, although strictly speaking, it is not the observation bias (defined against
68 the true mean state) that is corrected, but the lumped observation-bias-minus-forecast-bias.

69 The long data record of observed and forecast state estimates required for estimating
70 observation rescaling coefficients has slowed the implementation of land data assimilation
71 in large-scale applications, particularly within atmospheric systems, which are frequently
72 updated and yet prohibitively expensive to replay over long periods. Consequently, Dharssi
73 et al. (2011) and de Rosnay et al. (2013) identify the difficulty in obtaining observation
74 rescaling coefficients as one cause of the limited impact of assimilating remotely sensed soil
75 moisture observations into atmospheric models. The long data record requirement also pre-
76 vents the assimilation of new remotely sensed data sets, and necessitates costly reprocessing
77 of the rescaling parameters after significant updates to assimilated data sets.

78 Consequently, this manuscript presents a method for removing the O-F mean difference

79 (i.e., the lumped observation-bias-minus-forecast-bias) in land data assimilation systems
80 without access to a long data record, by using a two-stage observation bias and state update
81 estimation filter. ‘Bias’ is defined subjectively, in terms of the temporal and spatial scales
82 over which it applies. In seeking a bias correction method that does not require a long data
83 record, the bias is necessarily defined over shorter time scales, and the presented two-stage
84 filter dynamically estimates nonstationary O-F mean differences that evolve at seasonal time
85 scales.

86 There are typically large systematic differences between remotely sensed and modeled
87 T_{skin} (Ghent et al. 2010; Wang et al. 2014), and if not adequately addressed these differences
88 will result in a sub-optimal assimilation, potentially leading to degraded flux forecasts (e.g.,
89 Reichle et al. 2010). Hence, the two-stage observation bias and state estimation scheme has
90 been demonstrated here by assimilating geostationary T_{skin} observations into the Catchment
91 land surface model.

92 The remainder of this manuscript is outlined as follows. In Section 2, the two-stage
93 observation bias and state estimation scheme is developed, and contrasted to observation
94 rescaling approaches. The two-stage filter is then demonstrated with an example assimilation
95 of remotely sensed skin temperature (T_{skin}) observations into a land surface model. The T_{skin}
96 assimilation experiments are outlined in Section 3, before the results are presented in Section
97 4. Finally, Section 5 presents a summary and conclusions.

98 2. The state and bias filter equations

99 The two-stage observation bias and state estimation approach introduced here is based
100 on the on-line two-stage forecast bias and state estimation approach of Dee and Da Silva
101 (1998), which has been successfully implemented in atmosphere (Dee and Todling 2000),
102 ocean (Chepurin et al. 2005; Keppenne et al. 2005), and land (Bosilovich et al. 2007; De Lan-
103 noy et al. 2007; Reichle et al. 2010) data assimilation. Following Friedland (1969), Dee and

104 Da Silva (1998) decouple the forecast bias estimation from the state update, and use a sep-
105 arate Kalman filter to estimate the forecast bias. The (bias-blind) state update innovations
106 (i.e., the O-F residuals) are used to measure the forecast bias for the bias update, based on
107 the assumption that the observations are unbiased, and persistence is used to predict the
108 forecast bias. Pauwels et al. (2013) recently extended the theory of the two-stage forecast bias
109 and state estimation filter to also estimate the observation bias. In their approach, demon-
110 strated with synthetic experiments, the (bias-blind) state update innovation measures the
111 observation bias plus the forecast bias, and is partitioned into the two separate bias terms
112 by calibration. However, observations of the true mean state are ultimately required to
113 partition the sum of the biases.

114 In contrast, we derive the two stage filter as if to estimate the observation biases measured
115 using the (bias-blind) state update innovations, based on the assumption that the forecasts
116 are unbiased. However, in the intended land data assimilation applications, it is recognized
117 that the forecasts are almost certainly biased, so that the estimated ‘observation bias’ really
118 represents the O-F mean difference (the lumped observation-bias-minus-forecast-bias), to
119 be used to adjust the observations to have the same mean value as the forecast estimates,
120 consistent with observation rescaling approaches.

121 Below, the bias-free EnKF equations are reviewed (Section 2a), before the optimal so-
122 lution for the two-stage observation bias and state estimation filter is derived (Section 2b).
123 Then, a parameterization of the Kalman gain for the bias update is introduced, to avoid
124 specifying the unknown prior observation bias uncertainty (Section 2c).

125 *a. The bias-free EnKF*

126 The bias-free EnKF, as implemented by Reichle et al. (2013) for land data assimilation,
127 consists of a model forecast step and a state update step. For the i th ensemble member, the
128 state forecast and update at the k th assimilation time are:

$$x_{k,i}^- = f(x_{k-1,i}^+, q_{k,i}) \quad (1)$$

$$x_{k,i}^+ = x_{k,i}^- + K_k(y_{k,i}^o + H_k x_{k,i}^-) \quad (2)$$

$$y_{k,i}^o = y_k^o + v_{k,i} \quad (3)$$

where x is the model state vector, $f(\cdot)$ is the forecast model, q represents the model error (or perturbation vector), K is the Kalman gain matrix, y^o is the observation vector, H is the observation operator, and v is an applied (zero mean, normal) perturbation representative of the expected observation errors. For simplicity we assume H to be linear, however the theory is unchanged if this assumption is relaxed. Throughout this manuscript, a super-scripted state vector indicates an estimated value, with the $-$ and $+$ superscripts indicating the prior and posterior estimates, respectively. In contrast, the absence of a superscript for a state variable indicates the true state vector.

In a bias-free EnKF, the errors in x^- and y^o are assumed to have vanishing long-term mean errors, and to be uncorrelated with each other. Under these assumptions, x^+ provides an unbiased estimate of x , and the optimal (minimum posterior state error variance) Kalman gain for the k th state update, K_k , is given by:

$$K_k = P_k^{x-} H_k^T (R^o + H_k P_k^{x-} H_k^T)^{-1} \quad (4)$$

where P^{x-} is the prior model state error covariance matrix, and R^o is the observation error covariance matrix. P^{x-} is diagnosed from the ensemble spread, while for land data assimilation R^o is typically assumed to be constant in time and have zero off-diagonal terms (e.g., Draper et al. 2012). Applying the above equations in the presence of (unknown) observation and/or forecast biases is sub-optimal, and is referred to as ‘bias-blind’ data assimilation (Dee and Da Silva 1998).

149 *b. The two-stage observation bias and state estimation*

150 For an observation-bias-aware assimilation, the observation vector is allowed to have a
 151 nonzero mean error persisting over some extended time period (a bias). The biased obser-
 152 vations, written \tilde{y}_k^o , can be partitioned into the bias term, b_k , and the remaining zero-mean
 153 error component, y_k^o :

154

$$\tilde{y}_k^o = b_k + y_k^o \quad (5)$$

155 The observations are then bias-corrected within the state update (equation 2) to remove
 156 the bias from the innovations, giving an unbiased estimate of x^+ :

157

$$x_{k,i}^+ = x_{k,i}^- + \tilde{K}_k(\tilde{y}_{k,i}^o - b_k - H_k x_{k,i}^-) \quad (6)$$

158 where \tilde{K} is the Kalman gain for the state update based on the bias corrected observation
 159 vector.

160 A separate, discrete Kalman filter is then used to estimate the observation bias. The
 161 observation bias is measured using the mean O-F ($\langle \tilde{y}_{k,i}^o - H_k x_{k,i}^- \rangle$, where $\langle \cdot \rangle$ is the
 162 ensemble mean). The bias is initialized at zero, and persistence is used as the bias prediction
 163 model, since the bias is assumed not to change significantly during individual assimilation
 164 cycles. The persistence model is recognized as an approximation, since a (potentially desir-
 165 able) feature of the two-stage filter is the nonstationary nature of the bias estimates. The
 166 observation bias forecast and update equations for the k th assimilation time are then written:

167

$$b_k^- = b_{k-1}^+ \quad (7)$$

$$b_k^+ = b_k^- + L_k \langle \tilde{y}_{k,i}^o - b_k^- - H_k x_{k,i}^- \rangle \quad (8)$$

168 where L_k is the Kalman gain for the bias update. Equations 7 and 8 provide an unbiased
 169 estimate of the observation bias, regardless of the selection of L_k . Appendix A shows that if

170 the errors in the observations, the prior bias estimate, and the prior state estimate are not
 171 correlated with each other, and if b_k^- provides an unbiased estimate of the observation bias,
 172 the optimal (minimum error covariance) posterior bias estimate is obtained with L_k equal
 173 to:

174

$$L_k = P_k^{b-} (R^o + P_k^{b-} + H_k P_k^{x-} H_k^T)^{-1} \quad (9)$$

175 Here R^o is unchanged from equation 4 and represents the random errors in the observations
 176 only, while P_k^{b-} is the random error covariance matrix for the prior observation bias estimate.

177 Substituting the best estimate of the bias (b_k^+ ; equation 8) into equation 6 then gives the
 178 state update equation with observation bias correction:

179

$$x_{k,i}^+ = x_{k,i}^- + \tilde{K}_k (\tilde{y}_{k,i}^o - b_k^+ - H_k x_{k,i}^-) \quad (10)$$

180 Up to this point, the presented derivation of the two-stage observation bias and state
 181 estimation equations has followed that of Pauwels et al. (2013), with their forecast bias set
 182 to zero. However, we now diverge from their approach. In Appendix B, we show that if the
 183 optimal expression for L is used (equation 9), \tilde{K}_k in equation 10 is the same as K_k for the
 184 bias-free filter (equation 4). That is, the Kalman gain is unchanged by the inclusion of the
 185 two-stage observation bias estimate in the state update equation. This result parallels that
 186 of Dee and Todling (2000), who show that for the on-line two-stage forecast bias and state
 187 estimation filter the state update Kalman gain is unchanged by the inclusion of the forecast
 188 bias estimate in the state update equation.

189 To summarize the two-stage observation bias and state estimation filter equations pre-
 190 sented above, equations 1 and 10 are used for the state forecast and update, respectively,
 191 together with the state update Kalman gain of equation 4. Equations 7 and 8 are used
 192 for the observation bias forecast and update, respectively, together with the bias update
 193 Kalman gain of equation 9 (although equation 9 will be replaced by an empirical function in

194 Section c). For illustrative purposes, substituting equation 8 into equation 10, then taking
 195 the ensemble average gives:

196

$$x_{k,i}^+ = x_{k,i}^- + \tilde{K}_k(\tilde{y}_{k,i}^o - b_k^- - H_k x_{k,i}^-) - \tilde{K}_k L_k < \tilde{y}_{k,i}^o - b_k^- - H_k x_{k,i}^- > \quad (11)$$

197 and:

198

$$< x_{k,i}^+ > = < x_{k,i}^- > + \tilde{K}_k(I - L_k) < \tilde{y}_{k,i}^o - b_k^- - H_k x_{k,i}^- > \quad (12)$$

199 Comparing equation 12 to equation 8 for the bias update demonstrates how the two-stage
 200 filter partitions the innovations $(\tilde{y}_{k,i}^o - b_k^- - H_k x_{k,i}^-)$ into updates to the bias estimate and
 201 state estimate.

202 The presented two-stage observation bias and state estimation filter parallels the on-line
 203 two-stage forecast bias and state estimation of Dee and Da Silva (1998) but differs from the
 204 original two-stage estimation approach of Friedland (1969) in that the state update equation
 205 is optimized with the bias correction terms included (i.e., the Kalman gain is obtained by
 206 optimizing equation 10, rather than equation 2). The resulting two-stage filter is optimal if
 207 the various assumptions stated above hold. However, in practice the filter is unlikely to be
 208 optimal, since, for example, the prior state errors and the prior observation bias errors have
 209 been assumed uncorrelated, yet both contain information (and errors) from past observations.

210 *c. Parametrization of the bias gain*

211 The two-stage observation bias correction and state estimation approach outlined above
 212 requires the specification of the unknown error covariance matrix P^{b-} for the prior bias
 213 estimate to calculate the observation bias update Kalman gain, L , in equation 9. Dee
 214 and Da Silva (1998) and Pauwels et al. (2013) assumed that the prior forecast bias error
 215 covariances were proportional to the prior forecast error covariances, and Pauwels et al.

216 (2013) assumed that the prior observation bias error covariances were proportional to the
217 forecast observation error covariances. We instead replace L with an empirical function.
218 This approach is made possible because P^{b-} is not required for the bias-aware state update
219 Kalman gain, due to the equivalence of the bias-free and bias-aware Kalman gains noted in
220 Section b.

221 For the assimilation of a single observation type at a single location, L_k becomes scalar.
222 For the assimilation of the j th location and observation type, we approximate $L_{j,k}$ with a
223 function designed to approach one as the time since the last assimilated observation increases:

224

$$\lambda_{j,k} = 1 - e^{-\Delta t_{j,k}/\tau_j} \quad (13)$$

225 where $\Delta t_{j,k}$ is the number of time steps since the most recent observation of type j was
226 assimilated, and τ_j is a user-defined parameter representing the e-folding time scale of the
227 bias memory for observation type j . This function was chosen since it approximates the
228 expected behavior $L_{j,k}$ under two important scenarios. In the first scenario, no observations
229 have been recently assimilated, relative to the assumed time scale of the bias, and there is
230 little information with which to predict $b_{j,k}^-$. Hence, $L_{j,k}$ is expected to be close to one, as
231 predicted by equation 13 for large $\Delta t_{j,k}/\tau_j$. In the second scenario, observations are being
232 assimilated with some regularity, and random errors in $b_{j,k}^-$ will be dominated by random
233 errors in the $(\tilde{y}_k^o - H_k x_k^-)$ sequence used to update $b_{j,k}^-$ (since by definition the persistence
234 model will not introduce significant errors into the bias estimate), however, the bias filter
235 will gradually filter these errors over time. Hence, if $\Delta t_{j,k}$ is assumed to generalize the recent
236 availability of observations, equation 13 will approximate the increased certainty in $b_{j,k}^-$ (and
237 subsequent reduction in $\lambda_{j,k}$) as more observations are assimilated.

238 The empirical $\lambda_{j,k}$ must adequately account for the first scenario described above, of no
239 recent observations, since from equation 12 a large L_k is necessary in this case to prevent
240 the potentially large $b_{j,k}^-$ errors from being propagated into the model state vector. This
241 situation can occur reasonably regularly, since there are often seasonal-scale gaps in land

242 surface observation records, when atmospheric and/or land surface conditions prevent remote
243 sensing of the land surface. Note the contrast to forecast bias correction, for which one can
244 fall back on a conservative approach of underestimating the forecast bias (Dee and Todling
245 2000; Reichle et al. 2010) when the bias estimate is highly uncertain, since the model state
246 will still be updated towards the true state (defined by the observations in this case).

247 For the assimilation of multiple observation types and locations, $\lambda_{j,k}$ can be extended
248 in the obvious way to a matrix, Λ_k , by setting the j th diagonal element of Λ_k to $\lambda_{j,k}$,
249 and setting the off-diagonal terms to zero (i.e, disregarding potential spatial correlation, or
250 cross-correlation between observation types, in the bias updates). A potential weakness of
251 the above parameterization of $\lambda_{j,k}$ is that a $b_{j,k}^-$ estimate based on a single recent observation
252 would be assigned high confidence. Consequently, observations are excluded from the state
253 update when the bias estimate is based on less than two observations within the last $\tau_j/2$
254 time steps (although these observations are still used to update $b_{j,k}^-$).

255 *d. Comparison to observation rescaling*

256 The two-stage observation bias and state estimation method presented above treats the
257 systematic differences between observations and forecasts quite differently compared to the
258 observation rescaling methods currently used in many land data assimilation systems. Ob-
259 servation rescaling (Reichle and Koster 2004; Drusch et al. 2005; Scipal et al. 2008; Crow
260 et al. 2011) is designed to remove the long-term systematic differences in the mean and
261 variance (and possibly higher order moments) of the observed and forecast state estimates,
262 where ‘long-term’ is defined by the length of the data record used to calculate the rescal-
263 ing parameters. These systematic differences are typically assumed to be stationary, and a
264 static set of bias correction parameters is used. Consequently, a (bias-free) data assimilation
265 with observation rescaling will then adjust the model states to reduce residual differences
266 between the observations and model forecasts. Such differences include those occurring at
267 sub-seasonal time-scales, differences in the phase of the seasonal cycle, and also differences in

268 the intra-annual seasonal cycle, if the data record used to estimate the rescaling coefficients
269 was sufficiently long to sample the climatological inter-annual variability.

270 In contrast, the two-stage observation bias and state estimation method presented here
271 is designed to remove only the systematic difference in the mean of the observed and fore-
272 cast state estimates, and this mean difference is not restricted to being stationary. The
273 filter dynamically estimates the O-F mean differences based only on measurements up to the
274 current assimilation cycle, with greater weight placed on more recent measurements. The
275 resulting estimates are then nonstationary, and will evolve at a time scale determined by
276 the τ parameter in equation 13. Specifying τ to represent seasonal time scales will result in
277 the observations being adjusted to match the seasonal cycle of the forecast estimates. The
278 assimilation will then adjust the model state vector to reduce differences between the obser-
279 vations and forecasts at sub-seasonal time scales, somewhat consistent with the observation
280 rescaling approach. Although systematic differences in the variance of the observations and
281 forecasts are not explicitly removed, as they are in observation rescaling, the component of
282 variance due to seasonal, or longer, time scale dynamics will be addressed.

283 For a given data assimilation experiment, the suitability of the two-stage filter depends
284 on the distribution of the systematic differences between the observed and forecast esti-
285 mates. For T_{skin} , there can be large differences between the mean values of different model
286 forecast and observed estimates (Wang et al. 2014), however T_{skin} variability is reasonably
287 well constrained, due in part to the tight coupling between T_{skin} and the (comparatively well
288 observed) low-level atmospheric temperature. Hence, using the two-stage observation bias
289 and state estimation to adjust the seasonal cycle of the mean observed T_{skin} to match that
290 of the forecast estimates is expected to effectively address the systematic differences between
291 observed and forecast T_{skin} in an assimilation. However, for many other land surface vari-
292 ables this approach may not be sufficient. Most notably, for near-surface soil moisture there
293 are large systematic differences between the variability of different data sets, including the
294 sub-seasonal-scale variability (e.g., see Draper et al. (2013), their Figure 2). This is due in

295 part to the absence of global data sets constraining the possible soil moisture range, and
296 the subsequent uncertainty in the parameters controlling the soil moisture response to at-
297 mospheric forcing (specifically controlling the total volume of pore space available for water
298 storage in the soil column).

299 3. Skin temperature assimilation

300 The two-stage observation bias and state estimation scheme has been demonstrated by
301 assimilating geostationary T_{skin} observations into the Catchment land surface model. Two
302 separate assimilation experiments were performed. First, the T_{skin} data were assimilated
303 over the Americas at $0.3125^\circ \times 0.25^\circ$ longitude by latitude resolution, from 1 June, 2012 to
304 31 May, 2013. Second, to obtain example global maps of the mean differences between the
305 observed and forecast T_{skin} , the T_{skin} data were assimilated globally, at a coarser resolution
306 of $0.625^\circ \times 0.50^\circ$, from 1 May, to 1 August, 2012.

307 a. Catchment land surface model

308 Catchment (Koster et al. 2000) is the land surface modeling component of the Goddard
309 Earth Observing System Model, version 5 (GEOS-5; Rienecker et al. 2008). The catchment
310 model equivalent variable to remotely sensed T_{skin} is the surface temperature (T_{surf}), defined
311 as the average temperature of the canopy and soil surface, and representative of an arbitrarily
312 thin layer separating the canopy and soil surface from the atmosphere. While the Catchment
313 T_{surf} is prognostic, it has a very short memory over most land surface types due to its very
314 low surface specific heat capacity ($200 \text{ JK}^{-1}\text{m}^{-2}$, except for broadleaf evergreen vegetation).
315 The assimilation experiments were performed off-line (i.e., decoupled from the atmospheric
316 model), using meteorological forcing data from the NASA Modern-Era Retrospective analysis
317 for Research and Applications (MERRA) (Rienecker et al. 2011) and Catchment model
318 parameters from the routine GEOS-5 system. The initial land surface state was spun-up

319 from an archived GEOS-5 restart file on 1 January, 2000, by integrating the model forward
320 (without perturbations) to 1 January 2012, and the model ensemble was then spun up from
321 1 January, 2012 to the start of the assimilation on 1 June, 2012.

322 *b. Geostationary skin temperature data*

323 The assimilated T_{skin} observations are retrieved from geostationary Thermal Infrared
324 (TIR) brightness temperature observations at the NASA Langley Research Center (Scarino
325 et al. 2013). The T_{skin} data are retrieved every three hours, and reported on the $0.3125^{\circ} \times 0.25^{\circ}$
326 GEOS-5 model grid. The geostationary data have been produced in near-real time since
327 2011, from a constellation of satellites providing global (53° S to 53° N, after quality con-
328 trol) coverage: Geostationary Operational Environmental Satellites (GOES)-East, GOES-
329 West, the second Multifunctional Transport Satellite (MTSAT-2), Feng Yun-2E (FY-2E),
330 and Meteosat-9 (Met-9). However, for the assimilation experiment over the Americas do-
331 main, an updated data set from the GOES-East and GOES-West satellites, produced with
332 the latest retrieval model, has been used. Where observations are available from more than
333 one geostationary satellite, only the observations from the closest satellite were assimilated.
334 The observation quality control discards observations with a viewing zenith angle greater
335 than 60° , a solar zenith angle between 83° and 90° , a grid-cell cloud fraction above 20%, or
336 if the land modeling system indicates precipitation or a snow-covered surface.

337 Figure 1 shows the coverage of the observation-quality controlled (GOES-West and
338 GOES-East) T_{skin} observations assimilated in the Americas experiment, as a fraction of
339 the total number of possible observation times (eight 3-hourly observation times per day).
340 There are few observations available during colder periods, due mostly to increased cloudi-
341 ness. Hence, the coverage is very low (< 15% of the maximum possible coverage) at higher
342 latitudes. The coverage is also low over the Amazon, again due to cloudiness. There is some
343 diurnal variation in the coverage, with slightly more observations available during the day-
344 time hours (10% more than nighttime). In Section 4 evaluation statistics are only reported

345 at locations where observations were assimilated for at least 7.5% of the possible observation
346 times at each time of day (~ 30 observations).

347 *c. Assimilation system*

348 The state update component of the two-stage filter uses the EnKF (Reichle et al. 2013),
349 with 12 ensemble members and 3-hourly assimilation of the T_{skin} observations. The assimila-
350 tion update vector consists of T_{surf} and the ground heat content (GHT1) associated with
351 the near-surface (0-10 cm) soil temperature. The ensemble was generated using the forcing
352 and model state perturbations in Table 1, which were adapted from Reichle et al. (2010)
353 to account for the inclusion of GHT1 in the state update vector. Note that the Catchment
354 model version used in Reichle et al. (2010) had a much higher specific heat capacity for T_{surf}
355 ($70,000 \text{ JK}^{-1}\text{m}^{-2}$) than is currently used, and T_{surf} represented a 5 cm layer depth (hence
356 Reichle et al. (2010) updated only T_{surf}). The observation error standard deviations for the
357 T_{skin} retrievals were set at 1.3 K and 2.1 K during the nighttime and daytime, respectively,
358 which implies that the model and observations have roughly similar skill.

359 The Catchment model divides each model grid cell into multiple computational elements,
360 and a 3-D filter (with non-zero horizontal model and observation error correlations, Reichle
361 and Koster 2003) was used to spread the observations to all model computational surface
362 elements within each grid cell. For both the observation errors and the (forcing and state
363 vector) ensemble perturbations in Table 1, relatively short horizontal error correlation scales
364 with an e-folding distance of 0.17° were applied. Note that preliminary experiments with
365 increased horizontal error correlation scales (between 0.5° and 1.0°) degraded the assimila-
366 tion results, likely because the strong dependence on cloud cover limits the horizontal error
367 correlations of estimated T_{skin} .

368 The observation bias update was performed independently at each model grid cell (i.e.,
369 using a 1-D filter). Since there is a strong diurnal cycle in the observations-minus-forecast
370 mean difference (as will be shown in Section 4), the observation bias was modeled separately

371 at each of the eight diurnal observation times, following Reichle et al. (2010).

372 *d. Evaluation of assimilation output*

373 The results of the assimilation experiment over the Americas have been evaluated by
374 comparison to independent observations of clear sky T_{skin} , from the in situ SURFRAD net-
375 work (Augustine et al. 2005), and from remotely sensed MODIS TIR observations. The six
376 SURFRAD sites shown in Figure 1 were used (Fort Peck was excluded since the geostation-
377 ary satellite viewing zenith angle exceeds 60° there). For each of the validation sites, 3-hourly
378 T_{skin} were calculated from the SURFRAD up-welling and down-welling TIR radiation ob-
379 servations using the Stefan-Boltzmann equation, and broad-band emissivity calculated from
380 MODIS Terra monthly narrow-band emissivity observations (MOD11C3), using Wang et al.
381 (2005). For MODIS, Aqua (MYD11C1) and Terra (MOD11C1) daily clear-sky T_{skin} data
382 on the 0.05° Climate Modeling Grid have been averaged up to the GEOS-5 model grid,
383 and assumed to occur at the geostationary observation time closest to the median MODIS
384 observation time over the domain for each satellite orbit direction.

385 The skill of the T_{skin} assimilation experiment in predicting each of the independent data
386 sets has been compared to the skill of an open-loop (no data assimilation) ensemble, gener-
387 ated with the same model perturbations as used for the assimilation experiment. For both
388 cases, instantaneous model T_{surf} is compared to the independent T_{skin} observations at times
389 for which geostationary T_{skin} observations are available (i.e., for the assimilation experiment
390 the posterior T_{surf} is evaluated). There are systematic differences between the mean values
391 of the T_{skin} data sets used here, and these differences cannot be attributed to biases in any
392 particular data set. Hence, the evaluation statistic is the unbiased Root Mean Square Dif-
393 ference (ubRMSD), calculated at each model grid cell after removing the mean difference
394 over the full time period (separately at each time of day) between the data sets.

395 **4. Results**

396 *a. O-F mean differences*

397 Without bias correction there is a strong diurnal cycle in the mean difference between the
398 observed and forecast T_{skin} . For example, Figure 2 shows the diurnal cycle in the spatial mean
399 O-F mean difference over the Americas for a bias-blind assimilation of the GOES-East and
400 GOES-West geostationary T_{skin} observations (using the same observation error covariances
401 and forecast ensemble perturbations as for the bias-aware assimilation experiments). For
402 both GOES-East and GOES-West, the O-F mean differences are more positive after solar
403 noon. The GOES-West O-F mean differences are consistently positive, and larger than
404 those for GOES-East throughout the diurnal cycle, with a maximum value of 5.1 K at 21:00
405 UTC, compared to values < 2 K during the nighttime. In contrast, the GOES-East O-F
406 mean differences are negative during the nighttime, and positive during the daytime, but
407 with magnitude consistently < 1 K in both cases, except for the 2.8 K maximum at 18:00
408 UTC. The T_{skin} data retrieved from the different geostationary satellite are reasonably well
409 calibrated (Minnis et al. 2002), and the differences between the GOES-East and GOES-West
410 O-F mean differences in Figure 2 are almost certainly not related to the sensors themselves,
411 but to the contrasting land covers observed by each. The small spatial mean O-F mean
412 differences for GOES-East are due to cancellation between regions of positive and negative
413 O-F mean differences in the spatial means.

414 While the effectiveness of the observation bias correction has been analyzed throughout
415 the diurnal cycle, for brevity the focus here is on the results at 21:00 UTC, when the largest
416 O-F mean differences occurred in Figure 2. To demonstrate the influence of τ (the time scale
417 of the bias estimate) on the O-F mean differences estimated by the filter (i.e., the b^+), Figure
418 3 compares the b^+ time series at 21:00 UTC, estimated using τ values between 10 and 30 days,
419 at the three SURFRAD locations with the greatest observation coverage. The SURFRAD
420 locations are used only for convenience, and no SURFRAD data were used in these plots.

421 For comparison, each panel also includes a smoothed O-F time series, estimated using the
422 first two annual Fourier harmonics, following Vinnikov et al. (2008). Recall from Section 2d,
423 that selecting τ to represent seasonal time scales will allow the assimilation to correct for
424 sub-seasonal-scale (e.g., synoptic-scale) errors. The bias filter tracks the expected seasonal-
425 scale O-F mean differences, while filtering out the higher-frequency noise in the observed and
426 forecast T_{skin} . As expected, the filtered b^+ time series lag the smoothed time series, with the
427 lag increasing as τ increases in Figure 3. The minimum time scale of the features resolved by
428 the b^+ time series also increases as τ increases, and for shorter τ values there is more noise
429 around the seasonal cycle (particularly for 10 days). The greatest differences between the b^+
430 time series with different τ (and between the filtered and smoothed time series) occurred at
431 Sioux Falls, where the O-F seasonal cycle had the steepest temporal gradient. In particular,
432 during the 2012 summer when the O-F decreased rapidly, the b^+ time series are much higher
433 than the smoothed time series (likely due to the first two Fourier harmonics in the smoothed
434 time series being insufficient to capture the sharp gradient).

435 For a given application the best choice of τ for estimating the seasonal-scale O-F mean
436 differences will depend on the relative variability of the innovations at seasonal and sub-
437 seasonal time scales. For geostationary T_{skin} assimilation, a compromise value of $\tau = 20$
438 days has been selected, since this produced b^+ time series with reasonably smooth seasonal
439 cycles that did not lag the O-F time series by too much (Figure 3).

440 With a τ of 20 days, Figure 4 compares histograms of the state update innovations at
441 21:00 UTC at the same three locations plotted in Figure 3, for both the bias-blind assim-
442 ilation experiments and the two-stage observation bias and state estimation scheme. As
443 expected, the innovation distributions for the bias-blind assimilation are biased, with mean
444 values between 1.3 K and 8.0 K (Figures 4a-c). The inclusion of the observation bias correc-
445 tion reduced the mean innovations to magnitudes less than 0.5 K at each location (Figures
446 4d-f). The observation bias correction also changed the shape of the innovation distributions
447 in Figure 4, reducing their spread and skew. Consequently, the standard deviation at each

448 site is reduced, with the greatest reductions occurring at Sioux Falls, from 4.0 K for the
449 bias-blind assimilation to 2.5 K with the observation bias correction. The altered shape of
450 the innovation distribution is a consequence of the nonstationary bias estimation method
451 accounting for seasonal-scale evolution of the O-F mean difference. In contrast, if a single
452 (stationary) correction were applied to the mean over the full time period, the higher order
453 moments of the innovation distribution would have been unchanged.

454 The histograms in Figure 4 are representative of the performance of the observation bias
455 correction across the full domain, and throughout the diurnal cycle. For example, for both
456 satellites in Figure 2, the two-stage filter reduced the spatial mean O-F mean difference to
457 magnitudes between 0.0 - 0.3 K throughout the day, compared to bias-blind maxima of 5.1
458 K and 2.8 K, for GOES-West and GOES-East, respectively. Likewise the mean standard
459 deviation of the innovations across the domain was also reduced throughout the diurnal cycle
460 (not shown), for example from 3.8 K to 3.1 K for GOES-West, and from 2.7 K to 2.1 K for
461 GOES-East, both at 21:00 UTC.

462 Finally, in Section 2d it was hypothesized that for the assimilation of T_{skin} , the vari-
463 ability of modeled and observed estimates is reasonably well constrained so that adjusting
464 the mean seasonal cycle of the observations (with the two-stage filter) would be sufficient
465 to address the systematic differences between the observed and forecast estimates. Compar-
466 ing the variance of the observed and forecast T_{skin} confirms that this was the case in the
467 assimilation experiments performed here. For example, over the Americas at 21:00 UTC,
468 the spatially averaged temporal standard deviation of the GOES-West observations was 8.0
469 K, compared to 7.3 K for the model forecasts over the same domain, with a spatial mean
470 absolute difference between their standard deviations of 1.1 K. Likewise, for GOES-East at
471 21:00 UTC the mean standard deviation was 5.1 K, compared to 4.9 K for the forecasts,
472 with a spatial mean absolute difference of 0.9 K. The two-stage observation bias correction
473 reduced the differences in the variance, and the ‘bias corrected’ observations had spatially
474 averaged standard deviations very close to the model forecasts, of 7.6 K for GOES-West,

₄₇₅ with a spatial mean absolute difference of just 0.4 K, and of 5.1 K for GOES-East, giving a
₄₇₆ spatial mean absolute difference of 0.3 K.

₄₇₇ *b. Global O-F mean difference maps*

₄₇₈ Figure 5 shows maps of the estimated b^+ at 9:00 UTC on June 1, July 1, and August 1,
₄₇₉ 2012. There is substantial spatial variation in the b^+ , with a clear signal of land surface con-
₄₈₀ ditions. There are no obvious discontinuities between the b^+ estimated for adjacent satellites
₄₈₁ in Figure 5, although the limited regions of overlapping observations from neighboring satel-
₄₈₂ lites (at sufficiently small viewing angles) makes the direct assessment of such discontinuities
₄₈₃ difficult. At 9:00 UTC it is daytime over Africa and Europe, and this region has the largest
₄₈₄ estimated b^+ in Figure 5, with distinct regions of large positive values (> 10 K) in the drier
₄₈₅ regions of Africa, the Arabian peninsula, and western Asia, with a band of negative values
₄₈₆ (< -5 K) over equatorial Africa. In contrast, the regions experiencing nighttime generally
₄₈₇ have smaller b^+ (magnitude < 5 K), except for the drier regions of western North America
₄₈₈ and Australia, with mean differences of 5-10 K. This tendency for very large positive day-
₄₈₉ time b^+ over dry regions occurs consistently across the globe, particularly in the summer
₄₉₀ hemisphere; the same pattern was evident in Figure 2 for GOES-West, which observes the
₄₉₁ arid southwest of the US. In terms of the temporal evolution of the b^+ , the large-scale spatial
₄₉₂ patterns are consistent between the three months plotted in Figure 5, although the gradual
₄₉₃ evolution of the b^+ estimates is evident.

₄₉₄ *c. Evaluation against independent T_{skin} observations*

₄₉₅ Figures 6 and 7 demonstrate that the two-stage observation bias and state estimation
₄₉₆ filter improved the modeled T_{surf} sub-seasonal-scale variability, compared to independent
₄₉₇ observations, albeit by a modest amount. In Figure 6 the mean ubRMSD of the assimilation
₄₉₈ estimates versus SURFRAD observations is reduced at each time of day by between 0.05 K

499 - 0.31 K (\sim 5-10%), with the greatest improvements (>0.2 K) occurring during the first half
500 of the day (09:00-15:00 UTC). The ubRMSD across all times of day is significantly reduced
501 (at a 5% level) from 2.1 K to 1.9 K.

502 Similar results were obtained by comparison to Terra and Aqua MODIS T_{skin} observations
503 over the Americas, as shown in Figure 7. During the night, the open-loop ubRMSD was
504 already very small, with a spatial mean of 1.9 K for both Terra and Aqua. During the
505 day, the open-loop ubRMSD was much larger, except over the Amazon, with a spatial mean
506 of 3.6 K for both Terra and Aqua. For all MODIS overpasses, the assimilation consistently
507 improved the model fit to the MODIS data across the domain, except over the Amazon where
508 the open-loop ubRMSD was already very low and the improvement from the assimilation
509 was weaker, and even slightly negative in places. While the consistency of the positive
510 improvements in Figure 7 is encouraging, these improvements were significant (at the 5%
511 level) over only a small fraction ($<10\%$) of the domain. For each MODIS orbit direction
512 the spatial mean ubRMSD across the domain is shown in Table 2, and in each case the
513 assimilation reduced the spatial mean ubRMSD by around 10% of the open-loop value, with
514 ubRMSD reductions of 0.1 - 0.2 K during the nighttime, and 0.2-0.3 K during the daytime.

515 While the above evaluation consistently indicates that the T_{skin} assimilation has improved
516 the model T_{surf} , the improvements are rather modest. This is despite the use of only
517 assimilation update times in the evaluation, which will have exaggerated the assimilation
518 impact. There are several reasons for the modest results. Most importantly, the skill of the
519 model T_{surf} , in terms of the anomaly behavior assessed here, is already very good. Also, the
520 Catchment model T_{surf} has an extremely short memory, associated with its very low heat
521 capacity, hence the analysis updates do not persist into the subsequent model time step, and
522 the model has very little memory of improvements previously gained from the assimilation.
523 Including GHT1 in the state update vector did not increase the T_{surf} memory of previous
524 analysis updates, since the T_{surf} dynamics are dominated by the radiation budget. Finally,
525 the lack of memory is compounded by the low data volume associated with the lack of

526 TIR observations under cloudy conditions. The modest impact of the assimilation is not
527 related to the observation bias correction method, since similar results were obtained using
528 cumulative distribution functions (Reichle and Koster 2004) to rescale the observations (not
529 shown).

530 5. Summary and conclusions

531 A two-stage observation bias and state estimation scheme has been developed for use in
532 land data assimilation. In this scheme, the observation-minus-forecast (O-F) mean differ-
533 ences are estimated and removed from the innovations prior to updating the model state.
534 In applications where the model predictions are bias-free, the two-stage filter could also be
535 used to correct the observations towards the true mean state. The presented method is com-
536 putationally affordable, straightforward to implement in an existing assimilation, requires
537 specification of only a single additional parameter, and can be used to assimilate satellite
538 radiances or retrieved geophysical variables. In contrast to the observation rescaling meth-
539 ods currently used in land data assimilation systems, the two-stage filter does not require
540 a long data record. Hence, it has the potential to facilitate the use and success of land
541 data assimilation, particularly in atmospheric modeling systems for which long records of
542 consistently forecast land surface estimates are typically not available.

543 The two-stage filter includes a parameterization of the Kalman gain for the bias update
544 that introduces an explicit specification of the time scale of the O-F mean differences. Defin-
545 ing the O-F mean difference over seasonal time scales allows the assimilation to update the
546 model state vector in response to sub-seasonal-scale (e.g., synoptic scale) differences between
547 observed and forecast estimates.

548 In experiments assimilating geostationary T_{skin} observations into the Catchment land
549 surface model, the two-stage filter effectively removed the O-F mean difference from the
550 observations, and consequently improved synoptic-scale dynamics in the model T_{surf} (the

model equivalent variable to T_{skin}). These improvements were measured using the ubRMSD with independent estimates of T_{skin} from the SURFRAD network (at six sites in the US), and from MODIS satellite observations over the Americas. While modest, the improvements highlight the potential value of the geostationary T_{skin} for future modeling efforts.

Global maps of the O-F mean differences estimated by the two-stage filter show clear spatial coherence, with a signal of local land surface conditions. Most prominently, there is a strong tendency for large positive O-F differences in dry regions during the daytime. In this study, the O-F mean difference was estimated independently at each model grid cell. However, the spatial cohesion of the estimates suggests the potential to improve the two-stage filter design by incorporating horizontal information into the observation bias estimates. This could be achieved by either including spatial smoothing in the bias forecast model (assuming correlations between the O-F mean difference in adjacent areas), or by implementing the bias update using a 3-D filter (assuming correlations between the errors in the O-F mean difference estimates).

In addition to the difficulty of obtaining suitable data records for observation rescaling, several studies have highlighted other shortcomings arising from the stationary nature of the observation rescaling approaches for bias correction. For example, the inability of a stationary approach (CDF-matching) to distinguish between near-surface soil moisture variability over seasonal and sub-seasonal time scales can result in inadequate matching of the seasonal cycles between forecast estimates and CDF-matched observations (Draper et al. 2009). Also Drusch et al. (2005) argues that uncertainty in the inter-annual variability of the vegetation characteristics used in both soil moisture retrieval and land surface modeling may necessitate nonstationary observation bias correction methods, based on either frequent updates of observation rescaling coefficients, or the use of more sophisticated methods. More recently, Crow et al. (2011) showed that results from the assimilation of remotely sensed soil moisture into a simple water balance model were improved by using seasonally variable observation rescaling coefficients for adjusting the mean. The nonstationary nature of filtering may also

578 have practical advantages for the estimation of O-F mean differences, in that the estimates
579 can respond to step changes, caused for example, by changes in the forecast model, remote
580 sensor, or retrieval model. Hence, in atmospheric assimilation the ability of variational ob-
581 servation bias correction schemes to respond to temporal changes in the bias has proven
582 beneficial (Auligné et al. 2007; Dee and Uppala 2009).

583 Unlike observation rescaling, the two-stage filter presented here does not explicitly ad-
584 dress systematic differences between higher-order moments of the observations and the model
585 estimates. For the T_{skin} assimilation experiments presented here, the two-stage filter proved
586 sufficient. However, other land surface variables, including near-surface soil moisture, can
587 have large systematic differences in the sub-seasonal-scale variability of observed and forecast
588 estimates. Work is underway to expand the two-stage filter to also account for systematic dif-
589 ferences in the higher order moments, thus providing an alternative to observation rescaling
590 for soil moisture data assimilation.

591 *Acknowledgments.*

592 The research was supported by the NASA Modeling, Analysis, and Prediction program,
593 the NASA High-End Computing program, and the NASA Satellite Calibration Interconsis-
594 tency program. MODIS land surface data were provided by NASA's Earth Observing System
595 Data and Information System, and the SURFRAD data were provided by the NOAA Earth
596 System Research Laboratory.

APPENDIX

Appendix A. Derivation of L_k .

600 In the bias state update equation (equation 8), the model state, observation bias, and
 601 observation estimates can each be partitioned into their true value, a random (zero-mean)
 602 error, and for the observations a long term mean error (bias): $x_k^- = x_k + e^{x-}$, and $b_k^- =$
 603 $b_k + e^{b-}$, and $\tilde{y}_k^o = \tilde{y}_k + e_k^o = y_k + b_k + e_k^o$, where e represents the random error in the
 604 superscripted variable. To derive L_k , minimize the expected error in b_k^+ , $P_k^{b+} = E[e^{b+}(e^{b+})^T]$,
 605 where E is the expectation over time. Substituting equation 8 into P_k^{b+} , then partitioning
 606 each variable into its constituent parts gives:

$$P_k^{b+} = E[(b_k^+ - b_k)(b_k^+ - b_k)^T] \quad (A1)$$

$$= E[(b_k^- + L_k < \tilde{y}_k^o - b_k^- - H_k x_k^- > - b_k)(b_k^- + L_k < \tilde{y}_k^o - b_k^- - H_k x_k^- > - b_k)^T] \quad (A2)$$

$$= E[(e_k^{b-} + L_k < e_k^o - e_k^{b-} - H_k e_k^{x-} >)(e_k^{b-} + L_k < e_k^o - e_k^{b-} - H_k e_k^{x-} >)^T] \quad (A3)$$

608 The derivative of P_k^{b+} w.r.t L_k is:

$$\frac{\delta P_k^{b+}}{\delta L_k} = 2E[(e_k^{b-} + L_k < e_k^o - e_k^{b-} - H_k e_k^{x-} >)(< e_k^o - e_k^{b-} - H_k e_k^{x-} >)^T] \quad (A4)$$

610 Setting the derivative to zero, and solving for L , gives the P_k^{b+} minimum:

$$L_k = E[-e_k^{b-}(< e_k^o - e_k^{b-} - H_k e_k^{x-} >)^T(< e_k^o - e_k^{b-} - H_k e_k^{x-} >(< e_k^o - e_k^{b-} - H_k e_k^{x-} >)^T)^{-1}] \quad (A5)$$

612 If e_k^o , e_k^{b-} , and e_k^{x-} are not cross-correlated with each other, the expectation is:

$$L_k = P_k^{b-} (R^o + P_k^{b-} + H_k P_k^{x-} H_k^T)^{-1} \quad (A6)$$

615 Appendix B. Derivation of \tilde{K} , and equivalence to K .

616 To derive \tilde{K} minimize the expected error $x_{k,i}^+$, $P^{x+} = E[(e_k^{x+})(e_k^{x+})^T]$. Substituting
 617 equation 11 into P_k^{x+} , and as in Appendix A, partitioning each variable into its constituent
 618 terms, gives:

$$P^{x+} = E[(x_k^+ - x_k)(x_k^+ - x_k)^T] \quad (A7)$$

$$\begin{aligned} &= E[(x_k^- + \tilde{K}_k(\tilde{y}_k^o - b_k^- - H_k x_k^-) - \tilde{K}_k L_k < \tilde{y}_k^o - b_k^- - H_k x_k^- > - x_k) \\ &\quad (x_k^- + \tilde{K}_k(\tilde{y}_k^o - b_k^- - H_k x_k^-) - \tilde{K}_k L_k < \tilde{y}_k^o - b_k^- - H_k x_k^- > - x_k)^T] \end{aligned} \quad (A8)$$

$$\begin{aligned} &= E[(e_k^{x-} + \tilde{K}_k(e_k^o - e_k^{b-} - H_k e_k^{x-}) - \tilde{K}_k L_k < e_k^o - e_k^{b-} - H_k e_k^{x-} >) \\ &\quad (e_k^{x-} + \tilde{K}_k(e_k^o - e_k^{b-} - H_k e_k^{x-}) - \tilde{K}_k L_k < e_k^o - e_k^{b-} - H_k e_k^{x-} >)^T] \end{aligned} \quad (A9)$$

620 The derivative of P_k^{x+} w.r.t \tilde{K}_k is:

$$\begin{aligned} \frac{\delta P_k^{x+}}{\delta \tilde{K}_k} &= 2E[(e_k^{x-} + \tilde{K}_k(e_k^o - e_k^{b-} - H_k e_k^{x-}) - \tilde{K}_k L_k < e_k^o - e_k^{b-} - H_k e_k^{x-} >) \\ &\quad (e_k^o - e_k^{b-} - H_k e_k^{x-} - L_k < e_k^o - e_k^{b-} - H_k e_k^{x-} >)^T] \end{aligned} \quad (A10)$$

622 If e^o , e^{x-} , and e^{b-} are not cross-correlated with each other, setting the derivatives to zero
 623 to minimize P_k^{x+} , and taking the expectation gives:

$$\tilde{K}_k(I - L_k) = P_k^{x-} H_k^T (R^o + P_k^{b-} + H_k P_k^{x-} H_k^T)^{-1} \quad (A11)$$

625 Substituting equation 9 into A11 gives:

626

$$\tilde{K}_k(I - P_k^{b-}(R^o + P_k^{b-} + H_k P_k^{x-} H_k^T)^{-1}) = P_k^{x-} H_k^T (R^o + P_k^{b-} + H_k P_k^{x-} H_k^T)^{-1} \quad (A12)$$

$$\tilde{K}_k(R^o + P_k^{b-} + H_k P_k^{x-} H_k^T - P_k^{b-}) = P_k^{x-} H_k^T \quad (A13)$$

$$\tilde{K} = P_k^{x-} H_k^T (R^o + H_k P_k^{x-} H_k^T)^{-1} \quad (A14)$$

627 which is the same as equation 4 for the Kalman gain for the bias-free EnKF state update.
628 This demonstrates that the inclusion of the observation bias estimate from the two-stage
629 state and bias estimation does not change the expression of the solution for the Kalman gain
630 for the state update in equation 10 (assuming that equation 9 is used for L_k).

REFERENCES

633 Augustine, J., G. Hodges, C. Cornwall, J. Michalsky, and C. Medina, 2005: An up-
 634 date on SURFRAD: The GCOS Surface Radiation Budget Network for the continen-
 635 tal United States. *Journal of Atmospheric Oceanic Technology*, **22**, 1460–1472, doi:
 636 10.1175/JTECH1806.1.

637 Auligné, T., A. McNally, and D. Dee, 2007: Adaptive bias correction for satellite data
 638 in a numerical weather prediction system. *Q. J. R. Meteorol. Soc.*, **133**, 631–642, doi:
 639 10.1002/qj.56.

640 Bosilovich, M., J. Radakovich, A. da Silva, R. Todling, and F. Verter, 2007: Skin tempera-
 641 ture analysis and bias correction in a coupled land-atmosphere data assimilation system.
 642 *Journal of the Meteorological Society of Japan*, **85A**, 205–228, doi:10.2151/jmsj.85A.205.

643 Chepurin, G., J. Carton, and D. Dee, 2005: Forecast model bias correction in ocean data
 644 assimilation. *Mon. Wea. Rev.*, **133**, 1328–1342, doi:10.1175/MWR2920.1.

645 Crow, W., M. van den Berg, G. Huffman, and T. Pellarin, 2011: Correcting rainfall using
 646 satellite-based surface soil moisture retrievals: The Soil Moisture Analysis Rainfall Tool
 647 (SMART). *Water Resources Research*, **47**, W08521, doi:10.1029/2011WR010576.

648 Crow, W., et al., 2012: Upscaling sparse ground-based soil moisture observations for the
 649 validation of coarse-resolution satellite soil moisture products. *Reviews of Geophysics*, **50**,
 650 RG2002, doi:10.1029/2011RG000372.

651 De Lannoy, G., R. Reichle, K. Arsenault, P. Houser, S. Kumar, N. Verhoest, and V. Pauwels,
 652 2012: Multi-scale assimilation of AMSR-E snow water equivalent and MODIS snow cover
 653 fraction in northern Colorado. *Water Resources Research*, **48**, W01522.

654 De Lannoy, G., R. Reichle, P. Houser, V. Pauwels, and N. Verhoest, 2007: Correcting
655 for forecast bias in soil moisture assimilation with the ensemble Kalman filter. *Water*
656 *Resources Research*, **43**, W09410, doi:10.1029/2006WR005449.

657 de Rosnay, P., M. Drusch, D. Vasiljevic, G. Balsamo, C. Albergel, and L. Isaksen, 2013:
658 A simplified Extended Kalman Filter for the global operational soil moisture analysis at
659 ECMWF. *Quarterly Journal of the Royal Meteorological Society*, **139**, 1199–1213, doi:
660 10.1002/qj.2023.

661 Dee, D. and A. Da Silva, 1998: Data assimilation in the presence of forecast bias. *Q.J.R.*
662 *Meteorol. Soc.*, **124**, 269–295, doi:10.1002/qj.49712454512.

663 Dee, D. and R. Todling, 2000: Data assimilation in the presence of forecast bias: The
664 GEOS moisture analysis. *Mon. Wea. Rev.*, **128**, 3268–3282, doi:{10.1175/1520-0493(2000)
665 128<3268:DAITPO>2.0.CO;2}.

666 Dee, D. and S. Uppala, 2009: Variational bias correction of satellite radiance data in the
667 ERA-Interim reanalysis. *Q. J. R. Meteorol. Soc.*, **135**, 1830–1841, doi:10.1002/qj.493.

668 Dharssi, I., K. Bovis, B. Macpherson, and C. Jones, 2011: Operational assimilation of AS-
669 CAT surface soil wetness at the Met Office. *Hydrology and Earth System Sciences*, **15**,
670 2729–2746, doi:0.5194/hess-15-2729-2011.

671 Draper, C., J.-F. Mahfouf, and J. Walker, 2009: An EKF assimilation of AMSR-E soil mois-
672 ture into the ISBA land surface scheme. *Journal of Geophysical Research*, **114**, D20104,
673 doi:10.1029/2008JD011650.

674 Draper, C., R. Reichle, R. de Jeu, V. Naeimi, R. Parinussa, and W. Wagner, 2013: Esti-
675 mating root mean square errors in remotely sensed soil moisture over continental scale
676 domains. *Remote Sensing of Environment*, **137**, 288–298, doi:10.1016/j.rse.2013.06.013.

677 Draper, C., R. Reichle, G. De Lannoy, and Q. Liu, 2012: Assimilation of passive and active
678 microwave soil moisture retrievals. *Geophysical Research Letters*, **39**, L04401, doi:10.1029/
679 2011GL050655.

680 Drusch, M., E. Wood, and H. Gao, 2005: Observation operators for the direct assimilation
681 of TRMM microwave imager retrieved soil moisture. *Geophysical Research Letters*, **32**,
682 L15403, doi:10.1029/2005GL023623.

683 Fertig, E., et al., 2009: Observation bias correction with an ensemble Kalman filter. *Tellus*
684 *A*, **61**, 210–226, doi:10.1111/j.1600-0870.2008.00378.x.

685 Friedland, B., 1969: Treatment of bias in recursive filtering. *IEEE Transactions on Automatic*
686 *Control*, **14**, 359–367, doi:10.1109/TAC.1969.1099223.

687 Ghent, D., J. Kaduk, J. Remedios, J. Ardö, and H. Balzter, 2010: Assimilation of land
688 surface temperature into the land surface model JULES with an ensemble Kalman filter.
689 *Journal of Geophysical Research*, **115**, D19112, doi:10.1029/2010JD014392.

690 Harris, B. and G. Kelly, 2001: A satellite radiance-bias correction scheme for data assimila-
691 tion. *Q.J.R. Meteorol. Soc.*, **127**, 1453–1468, doi:10.1002/qj.49712757418.

692 Keppenne, C., M. Riener, N. Kurkowski, and D. Adamec, 2005: Ensemble Kalman filter
693 assimilation of temperature and altimeter data with bias correction and application to
694 seasonal prediction. *Nonlinear Processes in Geophysics*, **12**, 491–503.

695 Koster, R., M. Suarez, A. Ducharne, M. Stieglitz, and P. Kumar, 2000: A catchment-based
696 approach to modeling land surface processes in a general circulation model: 1. model struc-
697 ture. *Journal of Geophysical Research*, **105**, 24809–24822, doi:10.1029/2000JD900327.

698 Minnis, P., L. Nguyen, D. Doelling, D. Young, W. Miller, and D. Kratz, 2002: Rapid
699 Calibration of Operational and Research Meteorological Satellite Imagers. Part II: Com-

700 parison of Infrared Channels. *J. Atmos. Oceanic Technol.*, **19**, 1250–1266, doi:10.1175/1520-0426(2002)019<1250:RCOOAR>2.0.CO;2.

702 Pauwels, V., G. De Lannoy, H.-J. Hendricks Franssen, and H. Vereecken, 2013: Simul-
703 taneous estimation of model state variables and observation and forecast biases us-
704 ing a two-stage hybrid kalman filter. *Hydrol. Earth Syst. Sci.*, **17**, 3499–3521, doi:
705 10.5194/hess-17-3499-2013.

706 Reichle, R., G. De Lannoy, B. Forman, C. Draper, and Q. Liu, 2013: Connecting satellite
707 observations with water cycle variables through land data assimilation: Examples using the
708 NASA GEOS-5 LDAS. *Surveys in Geophysics*, **35**, 577–606, doi:0.1007/s10712-013-9220-8.

709 Reichle, R. and R. Koster, 2003: Assessing the impact of horizontal error correlations in
710 background fields on soil moisture estimation. *Journal of Hydrometeorology*, **4**, 1229–1242,
711 doi:10.1175/1525-7541(2003)004<1229:ATIOHE>2.0.CO;2.

712 Reichle, R. and R. Koster, 2004: Bias reduction in short records of satellite soil moisture.
713 *Geophysical Research Letters*, **31**, L19 501, doi:10.1029/2004GL020938.

714 Reichle, R., R. Koster, J. Dong, and A. Berg, 2004: Global soil moisture from satellite
715 observations, land surface models, and ground data: Implications for data assimila-
716 tion. *Journal of Hydrometeorology*, **5**, 430 – 442, doi:10.1175/1525-7541(2004)005<0430:
717 GSMFSO>2.0.CO;2.

718 Reichle, R., R. Koster, P. Liu, S. Mahanama, E. Njoku, and M. Owe, 2007: Comparison
719 and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning
720 Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel
721 Microwave Radiometer (SMMR). *Journal of Geophysical Research*, **112**, D09 108, doi:
722 {10.1029/2006JD008033}.

723 Reichle, R., S. Kumar, S. Mahanama, R. Koster, and Q. Liu, 2010: Assimilation of satellite-

724 derived skin temperature observations into land surface models. *Journal of Hydrometeo-*
725 *rology*, **11**, 1103–1122, doi:10.1175/2010JHM1262.1.

726 Rienecker, M., et al., 2008: The GEOS-5 Data Assimilation System - Documentation of
727 Versions 5.0.1, 5.1.0, and 5.2.0. *Technical Report Series on Global Modeling and Data*
728 *Assimilation*, **27**.

729 Rienecker, M., et al., 2011: MERRA - NASA's Modern-Era Retrospective Analy-
730 sis for Research and Applications. *Journal of Climate*, **24**, 3624–3648, doi:10.1175/
731 JCLI-D-11-00015.1.

732 Scarino, B., P. Minnis, R. Palikonda, R. Reichle, D. Morstad, C. Yost, B. Shan, and
733 Q. Liu, 2013: Retrieving clear-sky surface skin temperature for numerical weather pre-
734 diction applications from geostationary satellite data. *Remote Sensing*, **5**, 342–366, doi:
735 10.3390/rs5010342.

736 Scipal, K., M. Drusch, and W. Wagner, 2008: Assimilation of a ERS scatterometer derived
737 soil moisture index in the ECMWF numerical weather prediction system. *Advances in*
738 *Water Resources*, **31**, 1101–1112, doi:10.1016/j.advwatres.2008.04.013.

739 Slater, A. and M. Clark, 2006: Snow Data Assimilation via an Ensemble Kalman Filter. *J.*
740 *Hydrometeor.*, **7**, 478–493, doi:10.1175/JHM505.1.

741 Vinnikov, K., Y. Yu, M. Rama Varma Raja, D. Tarpley, and M. Goldberg, 2008: Diurnal-
742 seasonal and weather-related variations of land surface temperature observed from geosta-
743 tionary satellites. *Geophysical Research Letters*, **35**, L22708, doi:10.1029/2008gl035759.

744 Wang, A., M. Barlage, X. Zeng, and C. Draper, 2014: Comparison of land skin temperature
745 from a land model, remote sensing, and in-situ measurement. *Journal of Geophysical*
746 *Research*, **119**, doi:10.1002/2013JD021026.

747 Wang, K., Z. Wan, P. Wang, M. Sparrow, J. Liu, X. Zhou, and et al., 2005: Estimation of
748 surface long wave radiation and broadband emissivity using Moderate Resolution Imaging
749 Spectroradiometer (MODIS) land surface temperature/emissivity products. *Journal of*
750 *Geophysical Research*, **110**, D11109, doi:10.1029/2004JD005566.

751 **6. Figures**

₇₅₂ 7. Tables

753 **List of Tables**

754	1	Ensemble Generation Perturbation Parameters for Forcing and Model Prog-	
755		nostic Variables.	36
756	2	Spatial Mean of the ubRMSD (K) with MODIS T_{skin} Reported in Figure 7.	37

TABLE 1. Ensemble Generation Perturbation Parameters for Forcing and Model Prognostic Variables.

	(A)dditive, or (M)ultiplicative	Standard Deviation	AR(1) Time Scale	Perturbation cross-correlation			
				GHT1	T2m	SW	LW
T_{surf}	A	0.2 K	12 hours	0.7	0	0	0
GHT1	A	50,000 J	12 hours	-	0	0	0
2m air temp (T2m)	A	1 K	24 hours	-	0.4	0.4	
SW radiation	M	0.3	24 hours		-	-0.6	
LW radiation	A	20 Wm^{-2}	24 hours			-	

TABLE 2. Spatial Mean of the ubRMSD (K) with MODIS T_{skin} Reported in Figure 7.

Experiment	MODIS overpass			
	Nighttime		Daytime	
	Terra	Aqua	Terra	Aqua
Open-loop	1.89	1.94	3.62	3.60
T_{skin} assimilation	1.70	1.79	3.36	3.42
Difference	0.19	0.15	0.27	0.18

757 List of Figures

758 1 Coverage of the assimilated GOES-West and GOES-East T_{skin} observations
759 from 1 June, 2012 to 31 May, 2013, as a fraction of the maximum possi-
760 ble coverage (eight observations every day). The locations of the SURFRAD
761 measurement stations are marked as DRA (Desert Rock), TBL (Table Moun-
762 tain), SXF (Sioux Falls), GWN (Goodwin Creek), BON (Bondville), and PSU
763 (Penn State). The plotted meridians demark the GOES-West and GOES-East
764 domains. 40

765 2 Diurnal cycle of the T_{skin} O-F mean difference, averaged over the Americas, for
766 a bias-blind assimilation (solid) and the two-stage observation bias and state
767 estimation bias-aware assimilation with $\tau = 20$ days (dashed), for GOES-West
768 (black) and GOES-East (grey). 41

769 3 The T_{skin} O-F residuals [K] at 21:00 UTC (black crosses) at the a) Goodwin
770 Creek, b) Sioux Falls, and c) Desert Rock SURFRAD sites. Black lines show
771 the smoothed O-F time series using the first two annual Fourier harmonics.
772 Dots show the bias estimates from the two-stage observation bias correction
773 scheme using (dark blue) $\tau = 10$ days, (light blue) $\tau = 20$ days, and (pink)
774 $\tau = 30$ days. 42

775 4 Histograms of the state update innovations at 21:00 UTC, for the assimila-
776 tion of geostationary T_{skin} , at the Goodwin Creek (GWN), Sioux Falls (SXF),
777 and Desert Rock (DRA) SURFRAD sites, for a bias-blind assimilation (up-
778 per), and for the two-stage observation bias and state estimation bias-aware
779 assimilation with $\tau = 20$ days (lower). 43

780 5 Observation-minus-forecast T_{skin} mean difference, estimated at 09:00 UTC on
781 first a) June, b) July, and c) August, 2012. Values are shown only where the
782 observation bias estimate is considered valid for use in the state update equa-
783 tion. The plotted meridians demark the domain of each satellite: [-175°,-105°]
784 GOES-West, [-105°,-37°] GOES-East, [-37°, 54°] MTSAT-2, [54°,90°] FY-2E,
785 and [90°,-175°] Met-9.

44

786 6 ubRMSD with SURFRAD T_{skin} , calculated separately for each SURFRAD
787 site and each observation time, for the assimilation of geostationary obser-
788 vations with the two-stage filter (filled circles), and the open-loop (unfilled
789 circles). The mean ubRMSD at each time of day for the assimilation (open-
790 loop) is indicated by the solid (dashed) line.

45

791 7 ubRMSD with MODIS T_{skin} for the open-loop (upper), and the improvement
792 in the ubRMSD gained from the assimilating geostationary T_{skin} with the
793 two stage filter (lower: Δ ubRMSD=ubRMSD of open-loop - ubRMSD of
794 assimilation), separately for each Terra and Aqua overpass direction. Grey
795 indicates < 30 coincident geostationary and MODIS T_{skin} observations. The
796 plotted meridians demark the GOES-West and GOES-East domains.

46

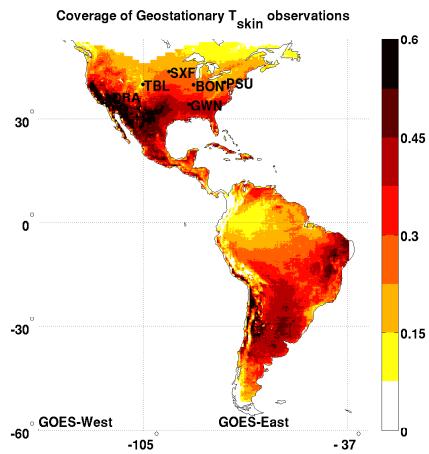


FIG. 1. Coverage of the assimilated GOES-West and GOES-East T_{skin} observations from 1 June, 2012 to 31 May, 2013, as a fraction of the maximum possible coverage (eight observations every day). The locations of the SURFRAD measurement stations are marked as DRA (Desert Rock), TBL (Table Mountain), SXF (Sioux Falls), GWN (Goodwin Creek), BON (Bondville), and PSU (Penn State). The plotted meridians demarcate the GOES-West and GOES-East domains.

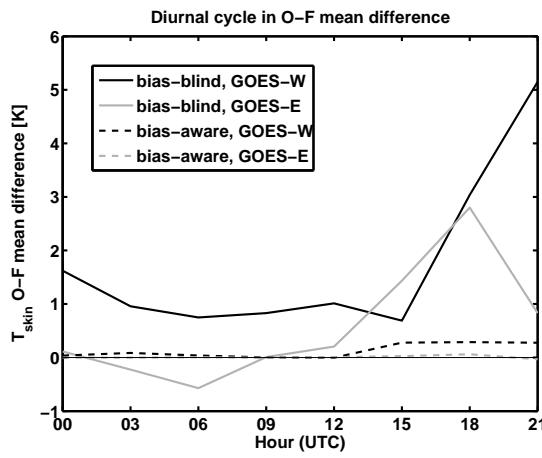


FIG. 2. Diurnal cycle of the T_{skin} O-F mean difference, averaged over the Americas, for a bias-blind assimilation (solid) and the two-stage observation bias and state estimation bias-aware assimilation with $\tau = 20$ days (dashed), for GOES-West (black) and GOES-East (grey).

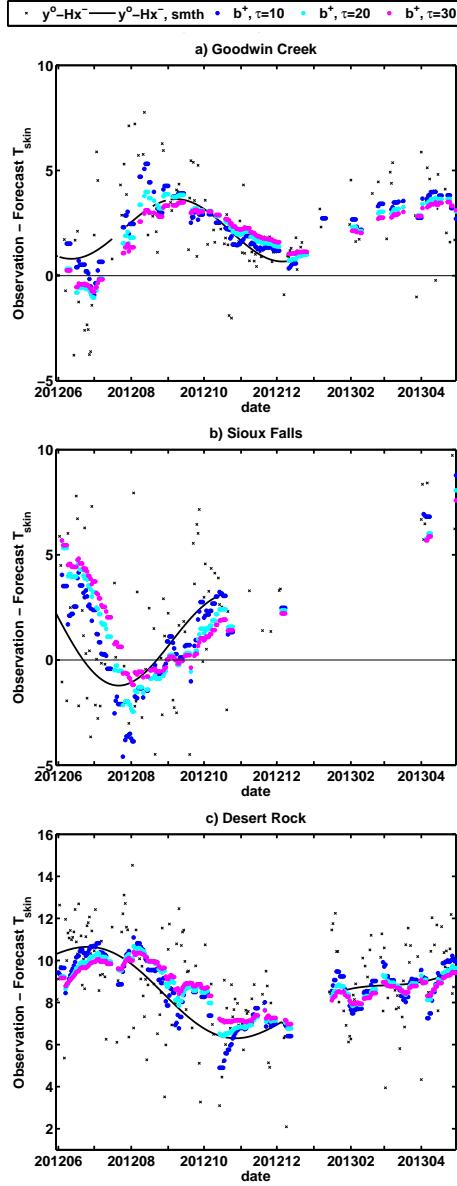


FIG. 3. The T_{skin} O-F residuals [K] at 21:00 UTC (black crosses) at the a) Goodwin Creek, b) Sioux Falls, and c) Desert Rock SURFRAD sites. Black lines show the smoothed O-F time series using the first two annual Fourier harmonics. Dots show the bias estimates from the two-stage observation bias correction scheme using (dark blue) $\tau=10$ days, (light blue) $\tau=20$ days, and (pink) $\tau=30$ days.

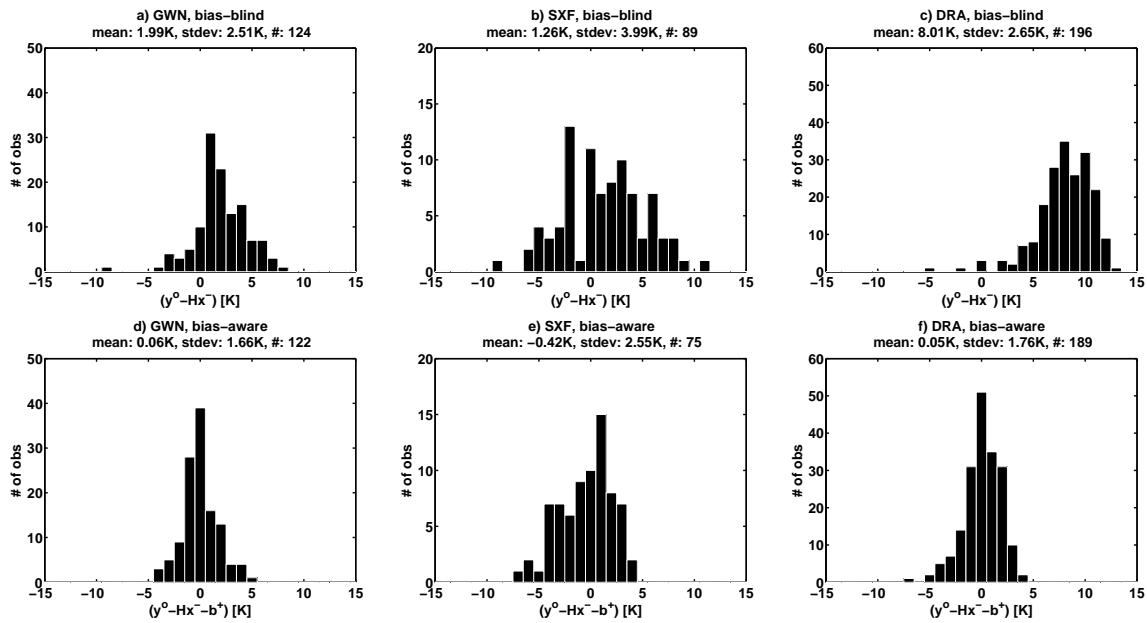


FIG. 4. Histograms of the state update innovations at 21:00 UTC, for the assimilation of geostationary T_{skin} , at the Goodwin Creek (GWN), Sioux Falls (SXF), and Desert Rock (DRA) SURFRAD sites, for a bias-blind assimilation (upper), and for the two-stage observation bias and state estimation bias-aware assimilation with $\tau=20$ days (lower).

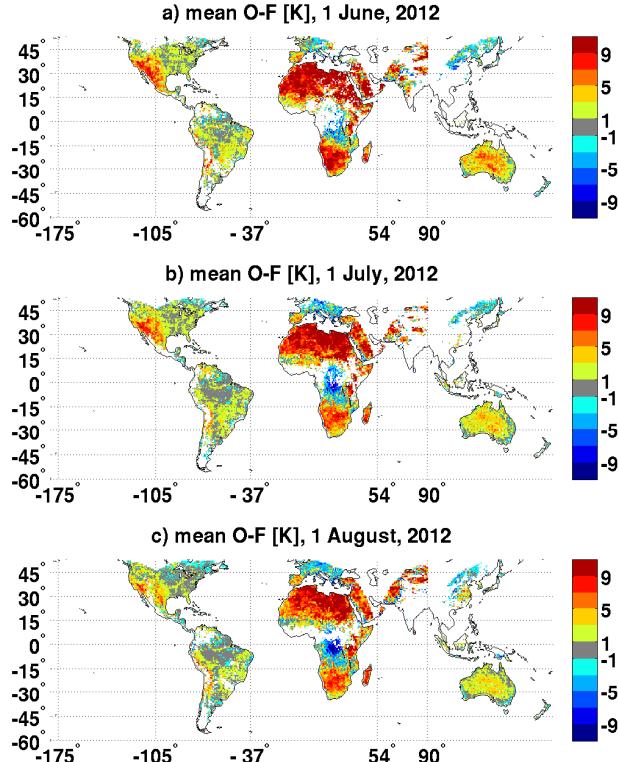


FIG. 5. Observation-minus-forecast T_{skin} mean difference, estimated at 09:00 UTC on first a) June, b) July, and c) August, 2012. Values are shown only where the observation bias estimate is considered valid for use in the state update equation. The plotted meridians demarcate the domain of each satellite: $[-175^\circ, -105^\circ]$ GOES-West, $[-105^\circ, -37^\circ]$ GOES-East, $[-37^\circ, 54^\circ]$ MTSAT-2, $[54^\circ, 90^\circ]$ FY-2E, and $[90^\circ, -175^\circ]$ Met-9.

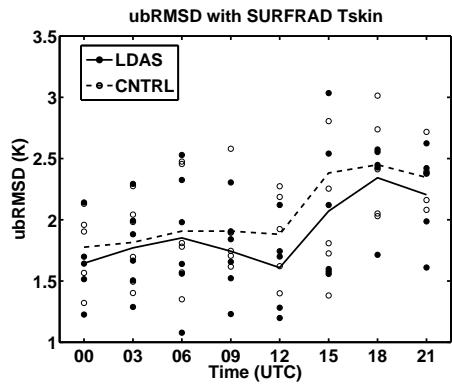


FIG. 6. ubRMSD with SURFRAD T_{skin} , calculated separately for each SURFRAD site and each observation time, for the assimilation of geostationary observations with the two-stage filter (filled circles), and the open-loop (unfilled circles). The mean ubRMSD at each time of day for the assimilation (open-loop) is indicated by the solid (dashed) line.

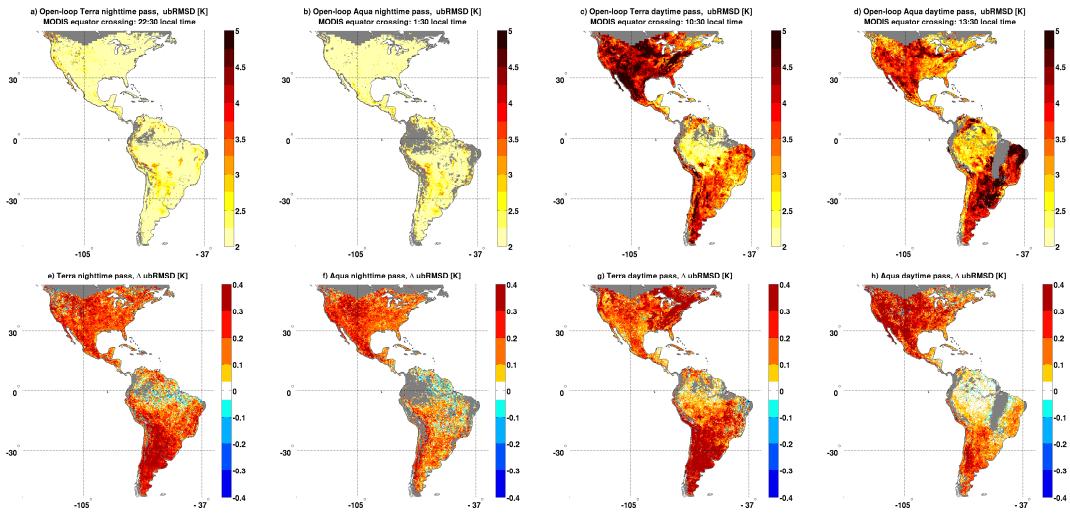


FIG. 7. ubRMSD with MODIS T_{skin} for the open-loop (upper), and the improvement in the ubRMSD gained from the assimilating geostationary T_{skin} with the two stage filter (lower: Δ ubRMSD=ubRMSD of open-loop - ubRMSD of assimilation), separately for each Terra and Aqua overpass direction. Grey indicates < 30 coincident geostationary and MODIS T_{skin} observations. The plotted meridians demarcate the GOES-West and GOES-East domains.